

ALGEBRAIC DEGREE OF POLYNOMIAL OPTIMIZATION

JIAWANG NIE AND KRISTIAN RANESTAD

ABSTRACT. Consider the polynomial optimization problem whose objective and constraints are all described by multivariate polynomials. Under some genericity assumptions, we prove that the optimality conditions always hold on optimizers, and the coordinates of optimizers are algebraic functions of the coefficients of the input polynomials. We also give a general formula for the algebraic degree of the optimal coordinates. The derivation of the algebraic degree is equivalent to counting the number of all complex critical points. As special cases, we obtain the algebraic degrees of quadratically constrained quadratic programming (QCQP), second order cone programming (SOCP) and p -th order cone programming (pOCP), in analogy to the algebraic degree of semidefinite programming [8].

1. INTRODUCTION

Consider optimization problem

$$(1.1) \quad \begin{cases} \min_{x \in \mathbb{R}^n} & f_0(x) \\ \text{s.t.} & f_i(x) = 0, i = 1, \dots, m_e \\ & f_i(x) \geq 0, i = m_e + 1, \dots, m \end{cases}$$

where $f_i(x)$ are multivariate polynomial functions in $\mathbb{R}[x]$ (the ring of polynomials in $x = (x_1, \dots, x_n)$ with real coefficients). The recent interest on solving polynomial optimization problems [6, 7, 10, 11] by using semidefinite relaxations or other algebraic methods motivates this study of the algebraic properties of the polynomial optimization problem (1.1). A fundamental problem about (1.1) is how the optimal solutions depend on the input polynomials $f_i(x)$. When the optimality condition holds and it has finitely many complex solutions, the optimal solutions are algebraic functions of the coefficients of polynomials $f_i(x)$, i.e., the coordinates of optimal solutions are roots of some univariate polynomials whose coefficients are functions of the input data. An interesting and important problem in optimization theory is to study the properties of these algebraic functions, e.g., how big their degrees are, i.e., what is the number of complex solutions to the critical equations of (1.1). Let us begin discussions with some special cases.

The simplest case of (1.1) is the linear programming (LP), i.e., all polynomials $f_i(x)$ have degree one. In this case, the problem (1.1) has the form (after removing the linear equality constraints)

$$(1.2) \quad \begin{cases} \min_{x \in \mathbb{R}^n} & c^T x \\ \text{s.t.} & Ax \geq b \end{cases}$$

where c, A, b are matrices or vectors of appropriate dimensions. The feasible set of (1.2) is now a polytope described by some linear inequalities. As is well-known, one optimal solution

Key words and phrases. algebraic degree, polynomial optimization, optimality condition, quadratically constrained quadratic programming (QCQP), p -th order cone programming, second order cone programming (SOCP), variety .

x^* (if it exists) of (1.2) must occur at one vertex of the polytope. So x^* can be determined by the linear system consisting of the active constraints. When the objective $c^T x$ is changing, the optimal solution might move from one vertex to another vertex. So the optimal solution is a piecewise linear fractional function of the input data (c, A, b) . When c, A, b are all rational, an optimal solution must also be rational, and hence its algebraic degree is one.

A more general convex optimization which is a proper generalization of linear programming is semidefinite programming (SDP) which has the standard form

$$(1.3) \quad \begin{cases} \min_{x \in \mathbb{R}^n} & c^T x \\ \text{s.t.} & A_0 + \sum_{i=1}^n x_i A_i \succeq 0 \end{cases}$$

where c is a constant vector and the A_i are constant symmetric matrices. The inequality $X \succeq 0$ means the matrix X is positive semidefinite. Recently, Nie et al. [8] studied the algebraic properties of semidefinite programming. When c and A_i are generic, the optimal solution x^* of (1.3) is shown [8] to be a piecewise algebraic function of c and A_i . Of course, the constraint of (1.3) can be replaced by the nonnegativity of all the principle minors of the constraint matrix, and hence (1.3) becomes a special case of (1.1). However, the problem (1.3) has very special nice properties, e.g., it is a convex program and the constraint matrix is linear with respect to x . Interestingly, if c and A_i are generic, the degree of each piece of this algebraic function only depends on the rank of the constraint matrix at the optimal solution. A formula for this degree is given in [8].

Another optimization problem frequently used in statistics and biology is the Maximum Likelihood Estimation (MLE), which has the standard form

$$(1.4) \quad \max_{x \in \Theta} p_1(x)^{u_1} p_2(x)^{u_2} \cdots p_n(x)^{u_n}$$

where Θ is an open subset of \mathbb{R}^n , the $p_i(x)$ are polynomials such that $\sum_i p_i(x) = 1$, and the u_i are given positive integers. The optimizer x^* is an algebraic function of (u_1, \dots, u_n) . This problem has recently been studied and a formula for the degree of this algebraic function has been found (cf. [1, 5]).

In this paper we consider the general optimization problem (1.1), when the polynomials f_0, f_1, \dots, f_m define a *complete intersection*, i.e., their common set of zeros has codimension $m + 1$. We show that an optimal solution is an algebraic function of the input data. We call the degree of this algebraic function the *algebraic degree* of the polynomial optimization problem (1.1). Equivalently, the algebraic degree equals the number of complex solutions to the critical equations of (1.1), when this is finite. Under some genericity assumptions, we give in this paper a formula for the algebraic degree of (1.1).

Throughout this paper, the words “generic” and “genericity” are frequently used. These words are given a precise meaning in algebraic geometry. Some property or condition holds “generically” means it holds in some Zariski open set (a set described by polynomial inequalities \neq). Any statement that is proved under such a genericity hypothesis will be valid for all data that lie in a dense, open subset of the space of data, and hence it will hold except on a set of Lebesgue measure zero.

The algebraic degree of polynomial optimization (1.1) addresses the computational complexity at a fundamental level. To solve (1.1) exactly essentially reduces to solving some univariate polynomial equations whose degrees are the algebraic degree of (1.1). As we can see later, the algebraic degree might be very big.

The paper is organized as follows. Section 2 derives a general formula for the algebraic degree, and Section 3 gives the formulae of the algebraic degrees for special cases like quadratically constrained quadratic programming, second order cone programming, and p -th order cone programming.

2. A GENERAL FORMULA FOR ALGEBRAIC DEGREE

In this section, we shall derive the formula for the algebraic degree of polynomial optimization problem (1.1), when the polynomials define a complete intersection. Suppose the polynomial $f_i(x)$ has degree d_i . Let x^* be one local or global optimal solution of (1.1).

At first, we assume all the inequality constraints are active, i.e., $m_e = m$, and the coefficients of polynomials f_1, f_2, \dots, f_m are generic. When $m = n$, by Bertini's Theorem [4, §17.16], the feasible set of (1.1) is finite and hence the algebraic degree is equal to the Bézout's number $d_1 d_2 \cdots d_m$. So, without loss of generality, assume $m < n$. If the variety

$$V = \{x \in \mathbb{C}^n : f_1(x) = \cdots = f_m(x) = 0\}$$

is smooth at x^* , i.e., the gradient vectors

$$\nabla f_1(x^*), \nabla f_2(x^*), \dots, \nabla f_m(x^*)$$

are linearly independent, then the *Karush-Kuhn-Tucker* (KKT) condition holds at x^* (Chapter 12 in [9]), i.e.,

$$(2.1) \quad \begin{cases} \nabla f_0(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0 \\ f_1(x^*) = \cdots = f_m(x^*) = 0 \end{cases}$$

where $\lambda_1^*, \dots, \lambda_m^*$ are Lagrange multipliers for constraints $f_1(x) = 0, \dots, f_m(x) = 0$. Thus the optimal solution x^* and Lagrange multipliers $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)$ are determined by the polynomial system (2.1). The set of points x^* in solutions to (2.1) forms the locus of critical points of (1.1). If the system (2.1) is zero-dimensional, then, by *elimination theory* [2], the coordinates of the points x^* are algebraic functions of the coefficients of the polynomials f_i . Each coordinate x_i^* can be determined by some univariate polynomial equation like

$$(x_i^*)^{\delta_i} + a_1(x_i^*)^{\delta_i-1} + \cdots + a_{\delta_i-1}x_i^* + a_{\delta_i} = 0$$

where a_j are rational functions of the coefficients of the f_i . Interestingly, when f_1, f_2, \dots, f_m are generic, the KKT condition always holds at any optimal solutions, and the degrees δ_i are equal to each other. This common degree counts the number of solutions to (2.1), i.e., the cardinality of the critical locus of (1.1) or, by definition, the algebraic degree of the polynomial optimization (1.1). We will derive a general formula for this degree.

In what follows, we work on the complex projective spaces, where the above question may be answered as a problem in intersection theory. For this we need to translate the optimization problem to a relevant intersection problem. Let \mathbb{P}^n be the n -dimensional complex projective space. A point $\tilde{x} \in \mathbb{P}^n$ is a class of vectors (x_0, x_1, \dots, x_n) that are parallel to each other. A variety in \mathbb{P}^n is a set of points \tilde{x} that satisfy a collection of homogeneous polynomial equations in (x_0, x_1, \dots, x_n) . Let $\tilde{f}_i(\tilde{x}) = x_0^{d_i} f_i(x/x_0)$ be the homogenization of $f_i(x)$. Define \mathcal{U} to be the projective variety in \mathbb{P}^n as

$$\mathcal{U} = \{\tilde{x} \in \mathbb{P}^n : \tilde{f}_1(\tilde{x}) = \tilde{f}_2(\tilde{x}) = \cdots = \tilde{f}_m(\tilde{x}) = 0\}.$$

Next, we let

$$\tilde{\nabla} \tilde{f}_i(\tilde{x}) = \left[\frac{\partial}{\partial x_0} \tilde{f}_i \quad \cdots \quad \frac{\partial}{\partial x_n} \tilde{f}_i \right]^T$$

be the gradient vector, with respect to the homogeneous coordinates. Notice that $(\frac{\partial}{\partial x_j} \tilde{f}_i = x_0^{d_i-1} \frac{\partial}{\partial x_j} f_i(x/x_0))$, so the homogenization of ∇f_i coincides with the last n coordinates in $\tilde{\nabla} \tilde{f}_i$.

In this homogeneous setting, the optimality condition for problem (1.1) with $m = m_e$ is

$$(2.2) \quad \left\{ (x, \mu) \in \mathbb{R}^n \times \mathbb{R} : \begin{array}{l} \tilde{f}_0(\tilde{x}) - \mu x_0^{d_0} = \tilde{f}_1(\tilde{x}) = \cdots = \tilde{f}_m(\tilde{x}) = 0 \\ \text{rank}[\tilde{\nabla}(\tilde{f}_0(\tilde{x}) + \mu x_0^{d_0}), \tilde{\nabla}(\tilde{f}_1(\tilde{x})), \dots, \tilde{\nabla}(\tilde{f}_m(\tilde{x}))] \leq m \end{array} \right\}$$

where $\mu \in \mathbb{R}$ is the critical value. Let $\tilde{x}^* \in \{x_0 \neq 0\}$ be a critical point, i.e., a solution to (2.2). We may eliminate μ by asking that the matrix

$$\begin{bmatrix} \tilde{f}_0(\tilde{x}^*) & \tilde{f}_1(\tilde{x}^*) & \cdots & \tilde{f}_m(\tilde{x}^*) \\ x_0^{d_0} & 0 & \cdots & 0 \end{bmatrix}$$

have rank one, and the matrix

$$\begin{bmatrix} \frac{\partial}{\partial x_0} \tilde{f}_0(\tilde{x}^*) & \frac{\partial}{\partial x_0} \tilde{f}_1(\tilde{x}^*) & \cdots & \frac{\partial}{\partial x_0} \tilde{f}_m(\tilde{x}^*) & (d_0 - 1)x_0^{d_0} \\ \frac{\partial}{\partial x_1} \tilde{f}_0(\tilde{x}^*) & \frac{\partial}{\partial x_1} \tilde{f}_1(\tilde{x}^*) & \cdots & \frac{\partial}{\partial x_1} \tilde{f}_m(\tilde{x}^*) & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial}{\partial x_n} \tilde{f}_0(\tilde{x}^*) & \frac{\partial}{\partial x_n} \tilde{f}_1(\tilde{x}^*) & \cdots & \frac{\partial}{\partial x_n} \tilde{f}_m(\tilde{x}^*) & 0 \end{bmatrix}$$

have rank $m + 1$. The first condition and the condition $x_0 \neq 0$ mean that our critical points

$$\tilde{x}^* \in \mathcal{U} = \{\tilde{f}_1(\tilde{x}) = \cdots = \tilde{f}_m(\tilde{x}) = 0\}$$

while the rank of the second matrix equals $m + 1$ at points where $x_0 \neq 0$ only if the submatrix

$$M = \begin{bmatrix} \frac{\partial}{\partial x_1} \tilde{f}_0(\tilde{x}) & \frac{\partial}{\partial x_1} \tilde{f}_1(\tilde{x}) & \cdots & \frac{\partial}{\partial x_1} \tilde{f}_m(\tilde{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_n} \tilde{f}_0(\tilde{x}) & \frac{\partial}{\partial x_n} \tilde{f}_1(\tilde{x}) & \cdots & \frac{\partial}{\partial x_n} \tilde{f}_m(\tilde{x}) \end{bmatrix}$$

has rank m . Therefore we define \mathcal{W} to be the projective variety in \mathbb{P}^n :

$$\mathcal{W} = \{\tilde{x} \in \mathbb{P}^n : \text{all the } (m+1) \times (m+1) \text{ minors of } M \text{ vanish}\},$$

the locus of points where the rank of $[\nabla(\tilde{f}_0), \dots, \nabla(\tilde{f}_m)]$ is less than or equal to m . Denote the class of $(1, x_1, \dots, x_n)$ in \mathbb{P}^n by \tilde{x} .

Proposition 2.1. *Assume $m = m_e$. If the polynomials f_1, \dots, f_m are generic, then we have:*

- (i) *The affine variety $V = \{x \in \mathbb{C}^n : f_1(x) = \cdots = f_m(x) = 0\}$ is smooth.*
- (ii) *The KKT condition holds at any optimal solution x^* ;*
- (iii) *If f_0 is also generic, the affine variety*

$$(2.3) \quad K = \left\{ x \in V : \exists \lambda_1, \dots, \lambda_m \text{ such that } \nabla f_0(x) + \sum_{i=1}^m \lambda_i \nabla f_i(x) = 0 \right\}$$

defined by KKT system (2.1) is finite.

Proof. (i) When polynomials f_1, \dots, f_m are generic, by Bertini's Theorem [4, §17.16], the variety \mathcal{U} has codimension m and is smooth, in particular the affine subvariety $V = \mathcal{U} \cap \{x_0 \neq 0\}$ is smooth. In terms of the Jacobian matrix

$$\begin{bmatrix} \frac{\partial}{\partial x_0} \tilde{f}_1(\tilde{x}) & \frac{\partial}{\partial x_0} \tilde{f}_2(\tilde{x}) & \cdots & \frac{\partial}{\partial x_0} \tilde{f}_m(\tilde{x}) \\ \frac{\partial}{\partial x_1} \tilde{f}_1(\tilde{x}) & \frac{\partial}{\partial x_1} \tilde{f}_2(\tilde{x}) & \cdots & \frac{\partial}{\partial x_1} \tilde{f}_m(\tilde{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_n} \tilde{f}_1(\tilde{x}) & \frac{\partial}{\partial x_n} \tilde{f}_2(\tilde{x}) & \cdots & \frac{\partial}{\partial x_n} \tilde{f}_m(\tilde{x}) \end{bmatrix},$$

its rank is full at \tilde{x} . Furthermore, the tangent space of V at \tilde{x} is, of course, not contained in the hyperplane $x_0 = 0$ at infinity, so the column $[1 \ 0 \ \cdots \ 0]^T$ is not in the column space of the matrix at \tilde{x} . Therefore already the submatrix

$$\begin{bmatrix} \frac{\partial}{\partial x_1} \tilde{f}_1(\tilde{x}) & \frac{\partial}{\partial x_1} \tilde{f}_2(\tilde{x}) & \cdots & \frac{\partial}{\partial x_1} \tilde{f}_m(\tilde{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_n} \tilde{f}_1(\tilde{x}) & \frac{\partial}{\partial x_n} \tilde{f}_2(\tilde{x}) & \cdots & \frac{\partial}{\partial x_n} \tilde{f}_m(\tilde{x}) \end{bmatrix}$$

has full rank at \tilde{x} , i.e., the gradients

$$\nabla \tilde{f}_1(\tilde{x}) \ \cdots \ \nabla \tilde{f}_m(\tilde{x})$$

are linearly independent at $\tilde{x} \in V$.

(ii) When x^* is one optimizer, which must belong to V , by (i), we know the gradients

$$\nabla f_1(x^*), \nabla f_2(x^*), \dots, \nabla f_m(x^*)$$

are linearly independent. Hence the KKT condition holds at x^* (Chapter 12 in [9]).

(iii) We claim that the intersection $\mathcal{U} \cap \mathcal{W}$ defined above is finite. Since our critical points $V \cap \mathcal{W}$ is a subset of $\mathcal{U} \cap \mathcal{W}$, (iii) would follow. The codimension of \mathcal{U} is m , and this variety is smooth, so the matrix M has by (i) rank at least m at each point of \mathcal{U} . The variety $\mathcal{U} \cap \{\tilde{f}_0(\tilde{x}) = 0\}$, is, by Bertini's Theorem, also smooth, so as above, the matrix M has full rank at points in the affine part $V \cap \{f_0(x) = 0\}$. On the other hand, M is the Jacobi matrix for the variety $\mathcal{U} \cap \{\tilde{f}_0(\tilde{x}) = 0\}$. This variety is again smooth and has codimension $m + 1$ in the hyperplane $\{x_0 = 0\}$, so M must have full rank $m + 1$ on $\mathcal{U} \cap \{\tilde{f}_0(\tilde{x}) = 0\}$. The variety \mathcal{W} where M has rank at most m , therefore cannot intersect $\mathcal{U} \cap \{\tilde{f}_0(\tilde{x}) = 0\}$. But Bézout's Theorem [3, §8.4] says that if the sum of the codimensions of two varieties in \mathbb{P}^n does not exceed n , then they intersect. In particular, any curve in \mathcal{U} intersects the hypersurface $\{\tilde{f}_0(\tilde{x}) = 0\}$. Since $\mathcal{U} \cap \{\tilde{f}_0(\tilde{x}) = 0\}$ has codimension $m + 1$, we deduce that \mathcal{W} must have codimension at least $n - m$. Furthermore, since any curve in $\mathcal{U} \cap \mathcal{W}$ would intersect $\{\tilde{f}_0(\tilde{x}) = 0\}$, the intersection $\mathcal{U} \cap \mathcal{W}$ must be empty or finite. On the other hand, the variety of $n \times (m + 1)$ -matrices with homogeneous forms as entries having rank no more than m has codimension at most $n - m$. So the codimension of \mathcal{W} equals $n - m$. Hence \mathcal{U} and \mathcal{W} have complementary dimensions. Therefore the intersection $\mathcal{U} \cap \mathcal{W}$ is non-empty and (iii) follows. \square

By Proposition 2.1, for generic f_1, \dots, f_m the optimal solutions of (1.1) can be characterized by the KKT system (2.1), and for generic objective function f_0 the KKT variety K is finite. Geometrically, the algebraic degree of the optimization problem (1.1) is, under these genericity assumption, equal to the number of distinct complex solutions of KKT, i.e., the cardinality of the variety K which we above showed to coincide with $V \cap \mathcal{W}$. The variety $\mathcal{U} \cap \mathcal{W}$ above clearly contains K . On the other hand, $\mathcal{U} \cap \mathcal{W}$ is finite and does not intersect

the hyperplane $\{x_0 = 0\}$ when polynomials f_i are generic. Since $\mathcal{U} - V = \mathcal{U} \cap \{x_0 = 0\}$ and the $\mathcal{U} \cap \mathcal{W} \cap \{x_0 = 0\} = \emptyset$, we can see that the cardinality of K coincides with the cardinality, i.e., the degree of $\mathcal{U} \cap \mathcal{W}$.

For integers (n_1, n_2, \dots, n_k) , define the symmetric sum of products as follows

$$(2.4) \quad D_r(n_1, n_2, \dots, n_k) = \sum_{i_1 + i_2 + \dots + i_k = r} n_1^{i_1} \cdots n_k^{i_k}.$$

Theorem 2.2. *Assume $m = m_e$. If the polynomials f_0, f_1, \dots, f_m are generic, then the algebraic degree of (1.1) is*

$$d_1 d_2 \cdots d_m D_{n-m}(d_0 - 1, d_1 - 1, \dots, d_m - 1).$$

Furthermore, if some f_i is not generic and the system (2.1) is zero-dimensional, then the above formula is an upper bound of the algebraic degree.

Proof. When f_1, f_2, \dots, f_m are generic, \mathcal{U} is a smooth complete intersection of codimension m . Its degree $\deg(\mathcal{U}) = d_1 d_2 \cdots d_m$. When f_0 is also generic, \mathcal{W} has codimension $n - m$ and intersects \mathcal{U} in a finite set of points as shown above. If the intersection $\mathcal{U} \cap \mathcal{W}$ is transverse (i.e., smooth) and hence consists of a collection of simple points, then the degree $\deg(\mathcal{U} \cap \mathcal{W})$ counts the number of intersection points of $\mathcal{U} \cap \mathcal{W}$, and hence the cardinality of KKT variety K , which is also the number of solutions to the KKT system (2.1) for problem (1.1).

To show that this intersection is transversal, we consider the subvariety X in $\mathbb{P}^n \times \mathbb{P}^m$ defined by the m equations $\tilde{f}_1 = \tilde{f}_2 = \dots = \tilde{f}_m = 0$ and the n equations

$$M \cdot (\lambda_0, \dots, \lambda_m)^T = 0,$$

where the λ_i are homogeneous coordinate functions in the second factor. The image under the projection of the variety X defined by these $m + n$ polynomials into the first factor coincides with the finite set $\mathcal{U} \cap \mathcal{W}$. Since M has rank at least m at every point of \mathcal{U} , there is a unique $\tilde{\lambda} = (\lambda_0, \dots, \lambda_m) \in \mathbb{P}^m$ for each point $\tilde{x} \in \mathcal{U} \cap \mathcal{W}$ such that $(\tilde{x}, \tilde{\lambda})$ lies in X . Therefore the X is a complete intersection. It is easy to check that this complete intersection does not have any fixed point when the coefficients of f_0 varies. So Bertini's Theorem [4, §17.16] applies to conclude that for generic f_0 this complete intersection is transversal, which implies that the intersection $\mathcal{U} \cap \mathcal{W}$ in \mathbb{P}^n is also transversal.

Since the intersection $\mathcal{U} \cap \mathcal{W}$ is finite, i.e., has codimension in \mathbb{P}^n equal to the sum of the codimensions of \mathcal{U} and \mathcal{W} , Bézout's Theorem (cf. [3, §8.4], [4, Theorem 18.3]) applies to compute the degree

$$\deg(\mathcal{U} \cap \mathcal{W}) = \deg(\mathcal{U}) \cdot \deg(\mathcal{W}).$$

To complete the computation, we therefore need to find $\deg(\mathcal{W})$. Since the codimension of \mathcal{W} equals the codimension of the variety defined by the $(m+1) \times (m+1)$ minors of a general $n \times (m+1)$ matrix with polynomial entries, the formula of Thom-Porteous-Giambelli [3, §14.4] applies to compute this degree: The degree of \mathcal{W} equals the degree of the determinantal variety of $n \times (m+1)$ matrices of rank at most m , in the space of matrices whose entries in the i -th column are generic forms of degree $d_i - 1$. These matrices may be considered as a collection of linear maps parameterized by \mathbb{P}^n . More precisely, they define a map between vector bundles of rank $m+1$ and $n+1$ over \mathbb{P}^n

$$\mathcal{M} : \mathcal{O}_{\mathbb{P}^n}(-d_0 + 1) \oplus \mathcal{O}_{\mathbb{P}^n}(-d_1 + 1) \oplus \cdots \oplus \mathcal{O}_{\mathbb{P}^n}(-d_m + 1) \longrightarrow \mathcal{O}_{\mathbb{P}^n}^{n+1},$$

and $\mathcal{W} \subset \mathbb{P}^n$ is the variety of points over which the map has rank m . The Thom-Porteous-Giambelli formula computes the degree in terms of the topological Chern classes of these

vector bundles: The degree equals the degree of the Chern class

$$c_{n-m}\left((\mathcal{O}_{\mathbb{P}^n}^{n+1}) - (\mathcal{O}_{\mathbb{P}^n}(-d_0+1) \oplus \mathcal{O}_{\mathbb{P}^n}(-d_1+1) \oplus \cdots \oplus \mathcal{O}_{\mathbb{P}^n}(-d_m+1))\right)$$

which coincides with the coefficient of t^{n-m} in

$$\frac{1}{(1 - (d_0 - 1)t) \cdot \dots \cdot (1 - (d_m - 1)t)} = (1 + (d_0 - 1)t + (d_0 - 1)^2 t^2 + \dots) \cdots (1 + (d_m - 1)t + (d_m - 1)^2 t^2 + \dots).$$

Thus $\deg(\mathcal{W})$ is the complete homogeneous symmetric function of degree $\text{codim}(\mathcal{W})$ evaluated at the column degree of G , which is $D_{n-m}(d_0 - 1, d_1 - 1, \dots, d_m - 1)$. Therefore the degree formula for the critical locus $\mathcal{U} \cap \mathcal{W}$ and hence the algebraic degree of (1.1) is proved.

When some polynomial f_i is not generic, then a perturbation argument can be applied. Let x^* be one fixed optimal solution of optimization problem (1.1). Apply a generic perturbation $\Delta_\epsilon f_i$ to each f_i so that $(f_i + \Delta_\epsilon f_i)(x)$ is a generic polynomial and the coefficients of $\Delta_\epsilon f_i$ tends to zero as $\epsilon \rightarrow 0$. Then one optimal solution $x^*(\epsilon)$ of the perturbed optimization problem (1.1) tends to x^* . By genericity of $(f_i + \Delta_\epsilon f_i)(x)$, we know

$$a_0(\epsilon)(x_i^*(\epsilon))^\delta + a_1(\epsilon)(x_i^*(\epsilon))^{\delta-1} + \cdots + a_{\delta-1}(\epsilon)x_i^*(\epsilon) + a_\delta(\epsilon) = 0.$$

Here $\delta = d_1 d_2 \cdots d_m D_{n-m}(d_0 - 1, d_1 - 1, \dots, d_m - 1)$ and $a_j(\epsilon)$ are rational functions of the coefficients of f_i and $\Delta_\epsilon f_i$. Without loss of generality, we can normalize $a_j(\epsilon)$ such that

$$\max_{0 \leq j \leq \delta} |a_j(\epsilon)| = 1.$$

When $\epsilon \rightarrow 0$, by continuity, we can see that x_i^* is a root of some univariate polynomial whose degree is at most δ and coefficients are rational functions of the coefficients of polynomials f_0, f_1, \dots, f_m . \square

Remark 2.3. The genericity assumption in the theorem is used to conclude that the critical locus $\mathcal{U} \cap \mathcal{W}$ is a smooth 0-dimensional variety by appealing to Bertini's Theorem [4, §17.16]. A sufficient condition for Bertini's Theorem to apply can be expressed in terms of the sets U_i of polynomials in which the polynomials f_0, f_1, \dots, f_m can be freely chosen. First assume that the generic polynomial in each U_i is reduced, and that U_i intersects every Zariski open set of a complex affine space V_i . Secondly, assume that the set of common zeros of all the polynomials in $\cup_{i=0}^m V_i$ is empty. Then Bertini's Theorem applies. In fact, the polynomials f_i for which the conclusion of Bertini's Theorem fails are contained in a complex subvariety of V_i .

If some of the polynomials f_i are reducible, then we may replace f_i by the factor of least degree that contains the optimizer. The original problem (1.1), is then modified to one with a smaller algebraic degree. This is relevant in the above context, if the generic polynomial in U_i is reducible.

Example 2.4. Consider the following special case of problem (1.1)

$$\begin{aligned} f_0(x) &= 21x_2^2 - 92x_1x_3^2 - 70x_2^2x_3 - 95x_1^4 - 47x_1x_3^3 + 51x_2^2x_3^2 + 47x_1^5 + 5x_1x_2^4 + 33x_3^5, \\ f_1(x) &= 88x_1 + 64x_1x_2 - 22x_1x_3 - 37x_2^2 + 68x_1x_2^2x_3 - 84x_2^4 + 80x_2^3x_3 + 23x_2^2x_3^2 - 20x_2x_3^3 - 7x_3^4, \\ f_2(x) &= 31 - 45x_1x_2 + 24x_1x_3 - 75x_3^2 + 16x_1^3 - 44x_1^2x_3 - 70x_1x_2^2 - 23x_1x_2x_3 - 67x_2^2x_3 - 97x_2x_3^2. \end{aligned}$$

Here $m = m_e = 2$. By Theorem 2.2, the algebraic degree of the optimal solution is bounded by

$$4 \cdot 3 \cdot D_1(4, 3, 2) = 12 \cdot (4 + 3 + 2) = 108.$$

Symbolic computation shows the optimal coordinate x_1 is a root of the univariate polynomial of degree 108 (whose coefficients are modulo 17)

$$\begin{aligned} & x_1^{108} + 8x_1^{107} + 7x_1^{106} + 4x_1^{105} - x_1^{104} - x_1^{103} + 2x_1^{102} - 7x_1^{100} - 7x_1^{99} + 7x_1^{98} + 5x_1^{95} - 4x_1^{94} - 6x_1^{93} + 4x_1^{92} - 8x_1^{91} + 6x_1^{90} \\ & + 4x_1^{89} + 6x_1^{88} + 2x_1^{87} + 6x_1^{86} - 7x_1^{85} - 3x_1^{84} + 5x_1^{83} - 6x_1^{82} - 3x_1^{81} + 8x_1^{80} - 4x_1^{79} - x_1^{78} - 2x_1^{77} + x_1^{76} - 3x_1^{75} + 6x_1^{74} \\ & + 7x_1^{73} + 4x_1^{72} + 3x_1^{71} - 4x_1^{70} - 8x_1^{68} - x_1^{67} - x_1^{66} + 2x_1^{65} + 6x_1^{64} - 4x_1^{63} + 5x_1^{62} + 2x_1^{61} + 4x_1^{60} - 2x_1^{59} - 5x_1^{58} + 7x_1^{57} \\ & - 8x_1^{56} + 5x_1^{55} + 8x_1^{54} - 8x_1^{53} - 2x_1^{52} - 2x_1^{51} - 4x_1^{50} - 3x_1^{49} + 5x_1^{48} - 6x_1^{46} + 6x_1^{45} - 6x_1^{44} + 5x_1^{43} + 5x_1^{42} - 5x_1^{41} - x_1^{40} \\ & + 5x_1^{39} - 4x_1^{38} - 3x_1^{37} + 5x_1^{36} - 2x_1^{35} - x_1^{34} - 6x_1^{33} - 8x_1^{32} + 6x_1^{31} + 6x_1^{30} + 8x_1^{29} + 4x_1^{28} - 8x_1^{27} - 5x_1^{26} - 4x_1^{25} + 2x_1^{24} \\ & - x_1^{23} + 2x_1^{22} + 3x_1^{21} + 2x_1^{20} + 4x_1^{19} + 6x_1^{18} + 5x_1^{17} - 7x_1^{16} - 2x_1^{15} - x_1^{14} - 7x_1^{13} + 5x_1^{12} + 2x_1^{11} - 8x_1^{10} - 5x_1^9 - 5x_1^8 \\ & - 3x_1^7 - 2x_1^5 - 7x_1^4 - 2x_1^3 - 6x_1^2 - 3x_1 - 1. \end{aligned}$$

In this case the degree bound 108 is sharp.

Now we consider the more general case that $m > m_e$, i.e., there are inequality constraints. Then a similar degree formula as in Theorem 2.2 can be obtained, when the active set is identified.

Corollary 2.5. *Let x^* be one optimizer and j_1, \dots, j_k be the active set of inequality constraints. If every active f_i is generic, then the algebraic degree of x^* is*

$$d_1 \cdots d_{m_e} d_{j_1} \cdots d_{j_k} D_{n-m_e-k}(d_0 - 1, d_1 - 1, \dots, d_{m_e} - 1, d_{j_1} - 1, \dots, d_{j_k} - 1).$$

If some f_i is not generic and the system (2.1) is zero-dimensional, then the above formula is an upper bound of the degree.

Proof. Note that x^* is also an optimal solution of polynomial optimization problem

$$\left. \begin{array}{ll} \min_{x \in \mathbb{R}^n} & f_0(x) \\ \text{s.t.} & f_i(x) = 0, i = 1, \dots, m_e \\ & f_i(x) = 0, i = j_1, \dots, j_k \end{array} \right\}.$$

Hence the conclusion follows from Theorem 2.2. \square

3. SOME SPECIAL CASES

In this section we derive the algebraic degree of some special polynomial optimization problems. The simplest special case is that all the polynomials f_i in (1.1) have degree one, i.e., (1.1) becomes one linear programming of the form (1.2). If the objective c is generic, precisely n constraints will be active. So the algebraic degree is $D_0(0, 0, \dots, 0) = 1$. This is consistent with we have observed in Introduction. Now let us look at other special cases.

3.1. Unconstrained optimization

We consider the special case that the problem (1.1) has no constraints. It becomes an unconstrained optimization. Its optimal solutions makes the gradient of the objective vanish. By Theorem 2.2, the algebraic degree is bounded by $D_n(d_0 - 1) = (d_0 - 1)^n$, which is exactly the Bézout's number of the gradient polynomial system

$$\nabla f_0(x) = 0.$$

Since f_0 can be chosen freely among all polynomials of degree d_0 , Remark 2.3 applies to show that the degree bound above is sharp.

Example 3.1. Consider the minimization of $f_0(x)$ given by

$$f_0 = x_1^4 + x_2^4 + x_3^4 + x_4^4 + x_1^3 + x_2^3 + x_3^3 + x_4^3 - 13x_1^2 - 30x_1x_2 - 9x_1x_3 + 5x_1x_4 + 11x_2^2 - 3x_3x_2 - 3x_3^2 - 20x_3x_4 - 13x_2x_4 - 9x_4^2 + x_1 - 2x_2 + 12x_3 - 13x_4.$$

For the above polynomial, the algebraic degree of the optimal solution is $3^4 = 81$. Symbolic computation shows the optimal coordinate x_1 of x^* is a root of the univariate polynomial of degree 81 (whose coefficients are modulo 17)

$$\begin{aligned} & x_1^{81} - x_1^{80} + 6x_1^{79} - x_1^{78} + 2x_1^{77} - 2x_1^{75} + 5x_1^{74} - 2x_1^{73} + 4x_1^{72} + 8x_1^{71} + 6x_1^{70} - x_1^{69} + 2x_1^{68} - 5x_1^{67} + 7x_1^{66} - 4x_1^{65} - 3x_1^{64} \\ & + 2x_1^{63} + 8x_1^{62} + 7x_1^{61} + 5x_1^{60} + 4x_1^{59} + 7x_1^{58} - 2x_1^{57} - 8x_1^{56} - 2x_1^{55} - 8x_1^{54} + 2x_1^{53} - 8x_1^{52} - x_1^{51} + 8x_1^{50} - 4x_1^{49} - 7x_1^{48} - x_1^{47} \\ & + 5x_1^{46} + 6x_1^{45} - 3x_1^{44} + 5x_1^{43} - 4x_1^{42} - 5x_1^{41} + x_1^{40} - 4x_1^{39} - 3x_1^{38} + 8x_1^{37} + 4x_1^{36} + 2x_1^{35} - 3x_1^{34} - 7x_1^{33} - 4x_1^{32} - 5x_1^{31} + 4x_1^{30} \\ & - 4x_1^{29} - 6x_1^{28} - 8x_1^{27} - 5x_1^{26} - 8x_1^{25} + 6x_1^{24} + 7x_1^{23} + 2x_1^{22} + 5x_1^{21} + x_1^{20} - 4x_1^{19} - 6x_1^{18} + 4x_1^{17} + 7x_1^{16} + 5x_1^{15} \\ & + 7x_1^{14} - 3x_1^{12} + 8x_1^{11} - x_1^{10} - 5x_1^9 - 4x_1^8 - 6x_1^7 + 7x_1^6 + 2x_1^5 + 3x_1^4 + 7x_1^3 - 3x_1^2 + 5x_1 - 6. \end{aligned}$$

The degree bound 81 is sharp for this problem.

3.2. Quadratic constrained quadratic programming

Consider the special case that all the polynomials f_0, f_1, \dots, f_m are quadratic. Then problem (1.1) becomes one quadratic constrained quadratic programming (QCQP) which has the standard form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & x^T A_0 x + b_0^T x + c_0 \\ \text{s.t.} \quad & x^T A_i x + b_i^T x + c_i \geq 0, \quad i = 1, \dots, \ell. \end{aligned}$$

Here A_i, b_i, c_i are matrices or vectors of appropriate dimensions. The objective and all the constraints are all quadratic polynomials. At one optimal solution, suppose $m \leq \ell$ constraints are active. By Corollary 2.5, the algebraic degree is bounded by

$$(3.1) \quad 2^m \cdot D_{n-m}(1, \underbrace{1, \dots, 1}_{m \text{ times}}) = 2^m \cdot \sum_{i_0 + i_1 + i_2 + \dots + i_m = n-m} 1 = 2^m \cdot \binom{n}{m}.$$

The polynomials f_0, f_1, \dots, f_m can be chosen freely in the space of quadratic polynomials, so Remark 2.3 applies to show that the degree bound above is sharp.

Example 3.2. Consider the polynomials

$$\begin{aligned} f_0 &= -20 - 27x_1^2 + 89x_1x_2 + 80x_1x_3 - 45x_1x_4 + 19x_1x_5 + 42x_1 - 13x_2^2 + 31x_2x_3 - 79x_2x_4 \\ &\quad + 74x_2x_5 - 9x_2 + 56x_3^2 - 77x_3x_4 - 2x_3x_5 + 35x_3 + 40x_4^2 - 13x_4x_5 + 60x_4 + 58x_5^2 - 84x_5, \\ f_1 &= 33 + 55x_1^2 - 41x_1x_2 + 33x_1x_3 - 61x_1x_4 + 96x_1x_5 + 12x_1 + 74x_2^2 - 90x_2x_3 - 57x_2x_4 \\ &\quad - 52x_2x_5 + 51x_2 + 15x_3^2 + 81x_3x_4 + 87x_3x_5 + 75x_3 - 10x_4^2 + 58x_4x_5 + 33x_4 + 83x_5^2 - 23x_5, \\ f_2 &= 8 - 9x_1^2 + 56x_1x_2 - 24x_1x_3 + 81x_1x_4 + 85x_1x_5 - 99x_1 - 77x_2^2 - 75x_2x_3 + x_2x_4 + 38x_2x_5 \\ &\quad + 23x_2 - 97x_3^2 - 14x_3x_4 - 73x_3x_5 + 65x_3 + 3x_4^2 - 14x_4x_5 + 16x_4 + 9x_5^2 - 10x_5, \\ f_3 &= 9 + 90x_1^2 - 94x_1x_2 - 22x_1x_3 - 24x_1x_4 + 78x_1 + 32x_2^2 - 48x_2x_3 - 6x_2x_4 + 80x_2x_5 - 18x_2 - 63x_3^2 \\ &\quad + 66x_3x_4 - 13x_3x_5 + 88x_3 - 45x_4^2 - 92x_4x_5 - 69x_4 - 43x_5^2 + 32x_5. \end{aligned}$$

For the above polynomials, the QCQP problem is nonconvex. We consider those local optimal solutions which make all the three inequalities active. By Corollary 2.5, the algebraic degree of

this problem is bounded by $2^m \binom{n}{m} = 80$. Symbolic computation shows the optimal coordinate x_1 is a root of the univariate polynomial of degree 80 (whose coefficients are modulo 17)

$$\begin{aligned} & x_1^{80} - 3x_1^{79} + 6x_1^{78} + 2x_1^{77} + 6x_1^{76} - 3x_1^{75} + 4x_1^{74} - 6x_1^{73} + x_1^{72} + 7x_1^{71} - 4x_1^{69} + 4x_1^{68} - 6x_1^{67} + x_1^{65} + 5x_1^{64} + x_1^{63} - 2x_1^{62} - 6x_1^{61} \\ & + 8x_1^{60} + 7x_1^{59} + x_1^{58} - 7x_1^{57} + 8x_1^{56} - 5x_1^{55} - x_1^{54} - 3x_1^{53} + x_1^{52} - 5x_1^{51} - 4x_1^{50} + 3x_1^{49} - 2x_1^{48} - x_1^{47} - 7x_1^{46} + 2x_1^{45} + 8x_1^{44} \\ & + 6x_1^{43} - 3x_1^{42} + 5x_1^{41} - 3x_1^{40} + 5x_1^{39} + 2x_1^{38} + 2x_1^{37} + 5x_1^{36} + x_1^{35} + 4x_1^{34} + 4x_1^{33} + x_1^{32} - x_1^{31} + 2x_1^{30} - x_1^{29} + 4x_1^{28} \\ & - 2x_1^{27} + 6x_1^{26} + 6x_1^{25} + 5x_1^{24} + 3x_1^{23} + 5x_1^{22} - x_1^{21} + 2x_1^{20} + 8x_1^{19} - x_1^{18} + 7x_1^{17} - x_1^{15} + x_1^{14} \\ & + 4x_1^{13} + 7x_1^{11} - 8x_1^{10} + 3x_1^9 + 6x_1^8 - x_1^7 + 8x_1^6 - 4x_1^5 + 8x_1^4 + x_1^3 - 2x_1^2 + 7x_1 - 4. \end{aligned}$$

The algebraic degree of this problem is 80 and the bound given by formula (3.1) is sharp.

3.3. Second order cone programming

The second order cone programming (SOCP) has the following standard form

$$(3.2) \quad \begin{aligned} & \min_{x \in \mathbb{R}^n} \quad c^T x \\ & s.t. \quad a_i^T x + b_i - \|C_i x + d_i\|_2 \geq 0, \quad i = 1, \dots, \ell \end{aligned}$$

where c, a_i, b_i, C_i, d_i are matrices or vectors of appropriate dimensions. Let x^* be one optimizer. Since SOCP is a convex program, the x^* must also be a global solution. By removing the square root in the constraint, SOCP becomes the polynomial optimization

$$\begin{aligned} & \min_{x \in \mathbb{R}^n} \quad c^T x \\ & s.t. \quad (a_i^T x + b_i)^2 - (C_i x + d_i)^T (C_i x + d_i) \geq 0, \quad i = 1, \dots, \ell. \end{aligned}$$

Without loss of generality, assume that the constraints with indices $1, 2, \dots, m$ are active at x^* . The objective is linear but the constraints are all quadratic. As we can see, the Hessian of the constraints has the special form $a_i a_i^T - C_i^T C_i$. Let r_i be the number of rows of C_i . When $r_i = 1$, the constraint $a_i^T x + b_i - \|C_i x + d_i\|_2 \geq 0$ is equivalent to two linear constraints

$$-(a_i^T x + b_i) \leq C_i x + d_i \leq a_i^T x + b_i.$$

Thus, when every $r_i = 1$, the problem reduces to a linear programming and hence has algebraic degree one, because in this situation the polynomial $(a_i^T x + b_i)^2 - (C_i x + d_i)^T (C_i x + d_i)$ is reducible. When $r_i \geq 2$ and a_i, b_i, C_i, d_i are generic, the polynomial $(a_i^T x + b_i)^2 - (C_i x + d_i)^T (C_i x + d_i)$ is quadratic of rank $r_i + 1$ and hence irreducible. Without loss of generality, assume $1 = r_1 = r_2 = \dots = r_k < r_{k+1} \leq \dots \leq r_m$. Then problem (3.2) is reduced to

$$\begin{aligned} & \min_{x \in \mathbb{R}^n} \quad c^T x \\ & s.t. \quad a_i^T x + b_i + \sigma_i (C_i x + d_i) \geq 0, \quad i = 1, \dots, k \\ & \quad (a_i^T x + b_i)^2 - (C_i x + d_i)^T (C_i x + d_i) \geq 0, \quad i = k+1, \dots, m \end{aligned}$$

where scalar σ_i is chosen such that $a_i^T x^* + b_i + \sigma_i (C_i x^* + d_i) = 0$. By Corollary 2.5, the algebraic degree of SOCP in this modified form is bounded by

$$(3.3) \quad 2^{m-k} \cdot D_{n-m}(0, \dots, 0, \underbrace{1, \dots, 1}_{m-k \text{ times}}) = 2^{m-k} \cdot \sum_{i_{k+1} + i_{k+2} + \dots + i_m = n-m} 1 = 2^{m-k} \cdot \binom{n-k-1}{m-k-1}.$$

When $k = m$, we have already seen the algebraic degree is one.

For the sharpness of degree bound (3.3), we apply Bertini's Theorem following Remark 2.3. For every $i = k+1, \dots, m$, define the set U_i of polynomials as

$$U_i = \left\{ (a_i^T x + b_i)^2 - \sum_{1 \leq j \leq r_i} \alpha_j^2 (C_i x + d_i)_j^2 : \alpha_1, \dots, \alpha_{r_i} \in \mathbb{R} \right\}.$$

Then define affine spaces V_i as follows:

$$V_i = \left\{ (a_i^T x + b_i)^2 - \sum_{1 \leq j \leq r_i} \beta_j (C_i x + d_i)_j^2 : \beta_1, \dots, \beta_{r_i} \in \mathbb{C} \right\}, i = k+1, \dots, m.$$

Then every set U_i intersects any Zariski open subset of the affine space V_i . On the other hand the set of common zeros of the linear polynomials

$$a_i^T x + b_i + \sigma_i (C_i x + d_i), i = 1, \dots, k$$

and all the polynomials in the union $\bigcup_{i=k+1}^m V_i$ is contained in the set

$$(3.4) \quad Z = \bigcap_{i=1}^k \left\{ x \in \mathbb{R}^n : a_i^T x + b_i + \sigma_i (C_i x + d_i) = 0 \right\} \bigcap_{i=k+1}^m \left\{ x \in \mathbb{R}^n : \begin{matrix} a_i^T x + b_i = 0 \\ C_i x + d_i = 0 \end{matrix} \right\}.$$

Therefore, for generic choices a_i, b_i, C_i, d_i , if $r_{k+1} + \dots + r_m + m > n$, the set Z is empty. Hence Remark 2.3 applies to show that, for generic choices of c, a_i, b_i, C_i, d_i , if $r_{k+1} + \dots + r_m + m > n$, the algebraic degree bound $2^{m-k} \cdot \binom{n-k-1}{m-k-1}$ is sharp.

Example 3.3. Consider SOCP defined by polynomials

$$\begin{aligned} f_0 &= -x_1 + 6x_2 + 13x_3 + 11x_4 + 8x_5, \\ f_1 &= (-11x_1 - 18x_2 - 4x_3 + 2x_4 - 12x_5 + 7)^2 - (-4x_1 - 10x_2 + 20x_3 - 4x_4 - 9x_5 + 3)^2 \\ &\quad - (-5x_1 - 11x_2 + 8x_3 - 18x_4 + 11x_5 + 15)^2 - (21x_1 + 18x_2 - 12x_3 - 10x_4 - 8x_5 + 4)^2, \\ f_2 &= (-5x_1 - 5x_2 - 7x_3 - 6x_4 + 4x_5 + 41)^2 - (x_1 - 2x_2 + 10x_3 - 21x_4 - 11)^2 \\ &\quad - (-12x_1 + 3x_2 + 16x_3 + 4x_4 + x_5 + 9)^2 - (14x_1 + 20x_2 - 13x_3 - 7x_4 + 4x_5 + 2)^2, \\ f_3 &= (x_1 - 8x_2 + 11x_3 - x_5 + 22)^2 - (2x_1 - x_2 + 3x_3 - x_4 - 25x_5 - 8)^2 \\ &\quad - (-2x_1 - 17x_3 + 14x_4 + 4x_5 - 7)^2 - (x_1 + 12x_2 + 14x_3 - 6x_4 - 4x_5 - 10)^2. \end{aligned}$$

There are no linear constraints. For this SOCP, all the three inequalities are active at the optimizer. All the matrices C_i has three rows. By formula (3.3), the algebraic degree of this problem is bounded by $2^3 \binom{5-1}{3-1} = 48$. Symbolic computation shows that the optimal coordinate x_1 is a root of the univariate polynomial of degree 48 (whose coefficients are modulo 17)

$$\begin{aligned} &x_1^{48} - 2x_1^{47} - 3x_1^{46} + 3x_1^{45} + 4x_1^{44} + 5x_1^{43} - 6x_1^{42} - 2x_1^{41} + 3x_1^{40} - 5x_1^{39} - 7x_1^{38} + 2x_1^{37} - 3x_1^{36} + 2x_1^{35} + 2x_1^{34} - 7x_1^{33} + 6x_1^{32} \\ &+ 3x_1^{31} + 3x_1^{29} + x_1^{28} - 6x_1^{27} - 3x_1^{26} + x_1^{25} + 4x_1^{24} - 7x_1^{23} - x_1^{22} + 5x_1^{21} + 3x_1^{20} - 4x_1^{19} + 2x_1^{18} - 8x_1^{17} + 5x_1^{16} \\ &+ 8x_1^{15} + 2x_1^{14} + 5x_1^{13} - 4x_1^{12} + 7x_1^{11} - 2x_1^{10} - 4x_1^9 + 4x_1^8 + 4x_1^7 + 3x_1^6 - 4x_1^5 - 5x_1^4 - 8x_1^3 + x_1^2 - x_1 - 1. \end{aligned}$$

The algebraic degree of this problem is 48, so the upper bound is sharp in this case.

3.4. p -th order cone programming

The p -th order cone programming (pOCP) has the standard form

$$(3.5) \quad \begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & a_i^T x + b_i - \|C_i x + d_i\|_p \geq 0, i = 1, \dots, \ell \end{aligned}$$

where c, a_i, b_i, C_i, d_i are matrices or vectors of appropriate dimensions. This is also a convex optimization problem. Let x^* be one optimizer, and assume the constraints with indices $1, \dots, m$ are active at x^* . Suppose the matrices C_i has r_i rows. When some $r_i = 1$, the constraint $a_i^T x + b_i - \|C_i x + d_i\|_p \geq 0$ is equivalent to two linear constraints

$$-(a_i^T x + b_i) \leq C_i x + d_i \leq a_i^T x + b_i.$$

Like the SOCP case, assume $1 = r_1 = \dots = r_k < r_{k+1} \leq \dots \leq r_m$. Then problem (3.5) is equivalent to

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & a_i^T x + b_i + \sigma_i(C_i x + d_i) \geq 0, \quad i = 1, \dots, k \\ & (a_i^T x + b_i)^p - \sum_{j=1}^{r_i} (C_i x + d_i)_j^p \geq 0, \quad i = k+1, \dots, m \end{aligned}$$

where σ_i is chosen such that $a_i^T x^* + b_i + \sigma_i(C_i x^* + d_i) = 0$. In this situation

$$D_{n-m}(0, \dots, 0, \underbrace{p-1, \dots, p-1}_{m-k \text{ times}}) = \sum_{i_{k+1} + \dots + i_m = n-m} (p-1)^{i_{k+1} + \dots + i_m} = (p-1)^{n-m} \binom{n-k-1}{m-k-1}.$$

By Corollary 2.5, the algebraic degree of x^* is therefore bounded by

$$(3.6) \quad p^{m-k} (p-1)^{n-m} \binom{n-k-1}{m-k-1}.$$

When $k = m$, problem (3.5) is reducible to some linear programming and hence its algebraic degree is one.

Now we discuss the sharpness of degree bound (3.6). Similarly to the SOCP case, for every $i = k+1, \dots, m$, define the set of polynomials U_i as

$$U_i = \left\{ (a_i^T x + b_i)^p - \sum_{1 \leq j \leq r_i} \alpha_j^p (C_i x + d_i)_j^p : \alpha_1, \dots, \alpha_{r_i} \in \mathbb{R} \right\}.$$

Then define affine spaces V_i as follows:

$$V_i = \left\{ (a_i^T x + b_i)^p - \sum_{1 \leq j \leq r_i} \beta_j (C_i x + d_i)_j^p : \beta_1, \dots, \beta_{r_i} \in \mathbb{C} \right\}, \quad i = k+1, \dots, m.$$

Then every set U_i intersects any Zariski open subset of the affine space V_i . On the other hand, the set of common zeros of the linear polynomials

$$a_i^T x + b_i + \sigma_i(C_i x + d_i), \quad i = 1, \dots, k$$

and all the polynomials in the union $\bigcup_{i=k+1}^m V_i$ is contained in the set Z defined by (3.4). Therefore, for generic choices of a_i, b_i, C_i, d_i , if $r_{k+1} + \dots + r_m + m > n$, the set Z is empty, and hence Remark 2.3 implies that the degree bound given by formula (3.6) is sharp.

Example 3.4. Consider the case $p = 4$ and the polynomials

$$\begin{aligned} f_0 &= 9x_1 - 5x_2 + 3x_3 + 2x_4 \\ f_1 &= (1 - 6x_1 - 6x_2 + 4x_3 - 9x_4)^4 - (7 - 6x_1 + 22x_2 - x_3 + x_4)^4 - (11 + x_1 - x_2 - 8x_3 + 3x_4)^4 \\ &\quad - (-13 + 7x_1 + 16x_2 - 7x_3 + 9x_4)^4 - (3 - 11x_1 + 14x_2 - 8x_3 + 5x_4)^4 - (8 + 9x_1 - 10x_2 + 2x_3 + 2x_4)^4. \end{aligned}$$

For the above polynomials, the inequality constraint must be active since the objective is linear. By formula (3.6), the algebraic degree of the optimal solution is bounded by $p^m(p-1)^{n-m}\binom{n-1}{m-1} = 108$. Symbolic computation shows the optimal coordinate x_1 is a root of the univariate polynomial of degree 108 (whose coefficients are modulo 17)

$$\begin{aligned} & x_1^{108} - 3x_1^{107} - 8x_1^{106} + 7x_1^{105} + 3x_1^{104} - 2x_1^{103} - 4x_1^{102} - 6x_1^{101} + 2x_1^{100} + 8x_1^{99} - 8x_1^{98} + 5x_1^{97} - 3x_1^{96} - 3x_1^{95} \\ & + 4x_1^{94} + 3x_1^{93} + 7x_1^{92} - 4x_1^{91} + 6x_1^{90} + x_1^{89} + 7x_1^{88} - x_1^{87} - 5x_1^{86} - 6x_1^{85} + x_1^{84} + 5x_1^{83} - x_1^{81} + 7x_1^{80} + 8x_1^{79} \\ & - 6x_1^{78} + 7x_1^{77} + 2x_1^{76} - 3x_1^{75} + 4x_1^{74} - 6x_1^{73} - 6x_1^{72} + x_1^{70} + 2x_1^{69} - x_1^{68} + 8x_1^{67} - 3x_1^{66} + 5x_1^{65} + 4x_1^{64} + x_1^{63} \\ & + x_1^{62} - 2x_1^{61} - x_1^{60} + 3x_1^{59} - 7x_1^{58} - 7x_1^{57} + 7x_1^{55} - 3x_1^{54} - 3x_1^{53} - 8x_1^{52} - 4x_1^{51} - 4x_1^{50} - 3x_1^{49} - 4x_1^{48} + x_1^{47} \\ & + 8x_1^{46} + 4x_1^{45} - 4x_1^{44} - 8x_1^{43} - 8x_1^{42} - 7x_1^{41} - 5x_1^{40} + 4x_1^{39} - 5x_1^{38} - 7x_1^{37} + 4x_1^{36} - 2x_1^{35} + x_1^{34} + 6x_1^{33} + 6x_1^{32} \\ & - 7x_1^{31} - 3x_1^{30} - 5x_1^{29} + 7x_1^{28} + 3x_1^{27} + 6x_1^{26} + 2x_1^{24} - 8x_1^{23} - 8x_1^{22} - 4x_1^{21} + 8x_1^{20} + 8x_1^{19} - 3x_1^{18} + 6x_1^{17} - 5x_1^{16} \\ & - 8x_1^{15} + 8x_1^{14} + 8x_1^{13} + 6x_1^{12} - 5x_1^{10} + 3x_1^9 + 2x_1^8 - 2x_1^7 + 6x_1^6 + 4x_1^5 + 7x_1^3 - 8x_1^2 + 4x_1 + 5. \end{aligned}$$

So the algebraic degree of this problem is 108, and the bound given by formula (3.6) is sharp.

Acknowledgement. The authors are grateful to Gabor Pataki, Richard Rimanyi and Bernd Sturmfels for motivating this paper.

REFERENCES

- [1] F. Catanese, S. Hosten, A. Khetan and B. Sturmfels: The maximum likelihood degree, *American Journal of Mathematics*, **128** (2006) 671–697.
- [2] D. Cox, J. Little and D. O’Shea. *Ideals, varieties, and algorithms. An introduction to computational algebraic geometry and commutative algebra*. Third edition. Undergraduate Texts in Mathematics. Springer, New York, 2007.
- [3] W. Fulton. *Intersection Theory*. Springer Verlag, 1984.
- [4] J. Harris. *Algebraic Geometry, A First Course*. Springer Verlag, 1992.
- [5] S. Hosten, A. Khetan and B. Sturmfels. Solving the likelihood equations. *Foundations of Computational Mathematics*. 5 (2005) 389–407.
- [6] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3): 796–817, 2001.
- [7] J. Nie, J. Demmel and B. Sturmfels. Minimizing polynomials via sum of squares over the gradient ideal. *Math. Prog.*, Series A, Vol. 106 (2006), No. 3, pp. 587–606.
- [8] J. Nie, K. Ranestad and B. Sturmfels. The algebraic degree of semidefinite programming. Preprint, 2006, [math.0611562](https://arxiv.org/abs/math/0611562).
- [9] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*, Springer Series in Operations Research, Springer-Verlag, New York, 1999.
- [10] P. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Math. Prog.*, Ser. B, Vol. 96, No.2, pp. 293–320, 2003.
- [11] P. A. Parrilo and B. Sturmfels. Minimizing polynomial functions. In S. Basu and L. Gonzalez-Vega, editors, *Algorithmic and Quantitative Aspects of Real Algebraic Geometry in Mathematics and Computer Science, volume 60 of DIMACS Series in Discrete Mathematics and Computer Science*, pages 83–99. AMS, 2003.

DEPARTMENT OF MATHEMATICS, UC SAN DIEGO, 9500 GILMAN DRIVE, LA JOLLA, CA 92093, USA
E-mail address: njw@math.ucsd.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OSLO, PB 1053 BLINDERN, 0316 OSLO, NORWAY
E-mail address: ranestad@math.uio.no